Date: _November 14, 2001_   Express Mail Label No. _EV005367542US_

Inventors: Vadim R. Viviani and Yoshihiro Ohmiya

Attorney's Docket No.:   2799.1001-002

NUCLEIC ACID MOLECULES ENCODING RED AND GREEN EMITTING LUCIFERASES

RELATED APPLICATIONS

This application is a continuation of U.S. Application No. 09/516,958, filed March 1, 2000, which is a continuation-in-part of U.S. Application No. 09/388,290, filed September 1, 1999, which claims priority to Italian application 5339, filed September 2, 1998, the contents of which are incorporated herein by reference in their entirety.

BACKGROUND OF THE INVENTION

Bioluminescence in beetles is characterized by a widerange of colors. Fireflies (Lampyridae) emit in the green-yellow region of the spectrum (*1, 2*), click-beetles (Elateridae) emit in the green-orange (*2, 3*), but railroad-worms (Phengodidae) span the widest range of the spectrum, that is, from the green to the red region (*4, 5*). The emission of green-red light was suggested to be an adaptation to optimize the detection of bioluminescence in distinct photic environments and for different biological functions (*6*). In all cases, such distinct colors arise from structurally homologous luciferases, which catalyze the same ATP-dependent oxidation of D-luciferin (*7*). Most studies about the structure and function of beetle luciferases have focused on a set of luciferases arising mainly from fireflies (*8-14*), two click-beetle species (*15, 16*), and recently a North American species of *Phengodes* (*17*). Three main factors at the level of

the luciferase active site are believed to govern bioluminescence colors (7): (I) the presence of basic residues catalyzing tautomerization between a ketonic (red light emitter) and enolic (yellow-green light emitter) species of excited oxy-luciferin (18-20); (II) the hydrophobicity of the active site (21, 22); and (III) the active site conformation which affects rotation of excited oxyluciferin along the C_2 -$C_{2'}$ bond (23). These factors may act together or independently to determine distinct bioluminescence colors in luciferases of different species. The construction of chimeric proteins using click-beetle luciferases (24) and firefly luciferases (25), along with mutagenesis studies (26-28), have revealed important regions and key residues for the bioluminescence color determination. The crystallographic structure of firefly luciferase has been recently resolved in the absence of the substrates (29), which shows a main N-terminal domain and a smaller C-terminal cleft which supposedly come closer to sandwich the substrates during catalysis. Despite all these studies, no structural investigations had been conducted on naturally occurring red light-emitting luciferases.

The beautiful and rare *Phrixothrix* railroad-worms are probably the most spectacular luminescent beetles, because in addition to their yellow-green bioluminescence (λ_{max} = 542-574 nm), displayed by two sets of 11 dorsal-lateral lanterns along the body, they emit red bioluminescence (λ_{max} = 609-638 nm) through cephalic and postcephalic organs (4, 5), a unique property among terrestrial creatures. The function of the lateral lantern bioluminescence is probably associated with defensive and sexual attraction purposes, whereas in the case of the red lantern bioluminescence was associated with self-illumination (5); however experimental evidence is still lacking. Due to their scarcity, only preliminary biochemical studies have been conducted about these creature luciferases (4, 20). Railroad-worm and click-beetle luciferases share a common feature: they do not suffer batchromic shift upon decreasing pH as lampyrid luciferases do (20). Due to their peculiar spectral properties, *Phrixothrix* luciferases constitute very important models for understanding the mechanism of color modulation in beetle bioluminescence.

SUMMARY OF THE INVENTION

The invention describes the cloning of the cDNAs arising from the beetles *Phrixothrix vivianii* and *Phrixothrix hirtus*, which code the luciferases that catalyze the production of green and red light, respectively. These cDNAs were characterized and

5   their respective sequences are shown in Figs. 1 and 2, along with their deduced amino acid sequences.

Thus, the present invention relates to an isolated nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of SEQ ID NO:1 and SEQ ID NO:3 and the complement of SEQ ID NO:1 and SEQ ID NO:3. The

10   invention further relates to a nucleic acid molecule which hybridizes under high stringency conditions to a nucleotide sequence selected from the group consisting of SEQ ID NO:1 and SEQ ID NO:3 and the complement of SEQ ID NO:1 and SEQ ID NO:3.

The invention also relates to a vector comprising an isolated nucleic acid

15   molecule of the invention operatively linked to a regulatory sequence, as well as to a recombinant host cell comprising the vector. The invention also provides a method for preparing a polypeptide encoded by an isolated nucleic acid molecule, comprising culturing a recombinant host cell (e.g., bacterial, fungal, plant, insect and mammalian cells) of the invention under conditions suitable for expression of said nucleic acid

20   molecule.

The invention further provides an isolated polypeptide encoded by isolated nucleic acid molecules of the invention. In a particular embodiment, the polypeptide comprises the amino acid sequence of SEQ ID NO: 2 or SEQ ID NO: 4. The invention also relates to an isolated polypeptide comprising an amino acid sequence which is

25   greater than about 80 percent identical and more specifically 90 percent identical to the amino acid sequence of SEQ ID NO: 2 or SEQ ID NO: 4.

The invention also relates to an antibody, or an antigen-binding fragment thereof, which selectively binds to the polypeptides of the invention, as well as to a method for assaying the presence of a polypeptide encoded by an isolated nucleic acid

molecule of the invention in a sample, comprising contacting said sample with an antibody which specifically binds to the encoded polypeptide.

BRIEF DESCRIPTION OF THE DRAWINGS

Figs. 1A-1C is a nucleotide sequence of the green light emitting luciferase cDNA of *Phrixothrix vivianii* (SEQ ID NO. 1) and the deduced amino acid primary structure (SEQ ID NO. 2).

Figs. 2A-2C is a nucleotide sequence of the red light emitting luciferase cDNA of *Phrixothrix hirtus* (SEQ ID NO. 3) and the deduced amino acid primary structure (SEQ ID NO. 4).

Fig. 3 is an *in vitro* bioluminescence spectra elicited by *Phrixothrix* railroad-worms recombinant luciferases: (A) Pv_{GR} and (B) Ph_{RE}. The BL reactions were conducted on 0.1 M Tris-HCl buffer, pH 8.0. These spectra were corrected for the spectral photosensitivity of the equipment and normalized.

The foregoing and other objects, features and advantages of the invention will be apparent from the following more particular description of preferred embodiments of the invention.

DETAILED DESCRIPTION OF THE INVENTION

Phrixothrix railroad-worms emit yellow-green light through 11 pairs of lateral lanterns along the body and red light through two cephalic lanterns. The cDNAs for the lateral lanterns luciferase of *Phrixothrix vivianii*, which emit green light (λ_{max} = 542 nm), and for the head lanterns of *P. hirtus*, which emit the most red-shifted bioluminescence (λ_{max} = 628 nm) among luminescent beetles, were cloned. Positive clones which emitted green (Pv_{GR}: λ_{max} = 549 nm) and red (Ph_{RE}: λ_{max} = 622 nm) bioluminescence were isolated. The lucifereases coded by Pv_{GR} (545 amino acid residues) and Ph_{RE} (546 amino acid residues) cDNAs share 71% identity. Pv_{GR} and Ph_{RE} luciferases showed 50-55% and 46-49% identity with firefly luciferases, respectively, and 47-49% with click-beetle luciferases. Ph_{RE} luciferase has some unique residues

which replace invariant residues in other beetle luciferases. The additional residue Arg 352 in Ph_{RE}, which is deleted in Pv_{GR} polypeptide, seems to be another important structural feature associated with red light production. As in the case of other railroad-worms and click-beetle luciferases studied, *Phrixothrix* luciferases do not undergo the typical red shift suffered by firefly luciferases upon decreasing pH, a property which might be related to the many amino acid residues shared in common between railroad-worm and click-beetle luciferase.

The complete cDNA nucleotidic sequences and the respective amino-acid sequences are shown in the universal genetic code. The region of the cDNA sequences which code for the luciferases (open reading frames) are shown in bold. The ATG start codons as well as the TAA stop codons are underlined.

In addition to the nucleotidic cDNA sequences, and the respective amino-acid sequences of the proteins, the two relevant properties that characterize the biological activity of these proteins (luciferases), are the catalysis of bioluminescence of distinct colors through the oxidation of the same substrate, the firefly D-luciferin [D-2-(6'hydroxy-2'benzothiaxolyl)Δ^2-thiazoline-4-carboxylic acid], as follows:

luciferase coded by *Phrixothrix vivianii* cDNA which produce green light (with maximum of electromagnetic spectrum centered at λ_{max}=549 nm)

luciferase coded by *Phrixothrix hirtus* cDNA which produce red light (with maximum of electromagnetic spectrum centered at λ_{max}=622 nm; other cloned luciferases emit in the range between λ_{max}=546-593 nm of the spectrum).

Comparison of Phrixothrix Luciferases.

Although a *Phengodes* luciferase had already been cloned, the sequences of *Phrixothrix* are the first ones of Phengodidae to be reported. The identity between *Phrixothrix* luciferases (71%) is lower than that expected for proteins of different species of the same genus, usually above 80% in the case of click-beetle (*16*) and of firefly (*11, 13*) luciferases. Some dissimilarity must result from the divergence of these luciferases toward the emission of distinct bioluminescence colors for the distinct

biological functions played by the lateral and head lanterns. Indeed, the luciferase of *R. ohbai* railroad-worm, which arises from lateral lanterns with function and bioluminescence colors similar to those of *P. vivianii* railroad-worm lateral lanterns (*37*), showed higher identity with Pv_{GR} luciferase (66%) than with Ph_{RE} luciferase (56%) that emits a color very different from that of a distinct lantern.

Comparison of *Phrixothrix* luciferases with a set of 12 other beetle luciferases showed 118 invariant residues, most of them located in the C-terminal region. Beside these invariant residues, *Phrixothrix* luciferases showed 21 additional residues in common with click-beetle luciferases and 31 residues with firefly luciferases. Most of the residues in common with click-beetle luciferases are located in the region between residues 242 and 333, whereas those in common with firefly luciferases are preferentially located in the region between residues 310 and 350. The peroxisomal targeting tripeptide SKL was found in all railroad-worm and click-beetle luciferases and in most firefly luciferases.

Residues Typical of Ph_{RE} Luciferase.

The region from residue 300 to about 480 appears to be more homologous with Pv_{GR} and *R. ohbai* green light-emitting luciferases than with Ph_{RE} luciferase. In Ph_{RE} luciferase, the occurrence of many unique substitutions in this otherwise conserved region suggests that this region plays some role in the bioluminescence color determination. The substitution of Ala314 by Ser in Ph_{RE} luciferase is located in a fragment which supposedly interacts with luciferin according to a recently proposed model (*42*). Whereas substitution of Ala 314 by Ser is associated with considerable chemical changes, the substitution of Ile410 by Leu (Ph_{RE} numbering), except for sterical hindrance effects, does not seem to have considerable influence since these residues have very similar chemical properties. Other substitutions lie in the region 430-480, mainly in the fragment 469-479 which displays a quite hydrophilic character in relation to other luciferases. In click-beetle isoenzymes, which emit such different colors as green and orange, this fragment is invariable and thus does not account for

color determination. If such substitutions influence bioluminescence spectra through a cumulative effect, it is more likely that they do so through influencing the conformation of active site, rather than through a solvent effect, since no trend indicative of a relationship between hydropathy profiles and bioluminescence colors was found.

5   Furthermore, the solvent effect created by many substitutions on the active site, although influencing short-range spectral shifts (<40 nm), is not enough only by itself to explain green-red shift, since fluorescence studies on oxyluciferin, dehydroluciferin, and analogues, in solvents with distinct dielectric constants, failed to get such a large spectral shift as that observed in *Phrixothrix* luciferases (>70 nm) (*45, 46*).

10   The presence of Arg352 in Ph_{RE} luciferase, which corresponds to the deleted residue in the shorter Pv_{GR} luciferase, is another important distinctive structural feature between these proteins. In Ph_{RE} this region (residues 350-362) is associated with a large increase of hydrophobicity in relation to Pv_{GR} luciferase. Such a feature could be potentially involved with considerable conformational changes among these luciferases.

15 Such changes could affect the bioluminescence colors through proper positioning a basic residue in the neighbors of C-5 of excited oxyluciferin in the case of yellow-green-emitting luciferases (*20*), or by influencing the active site geometry according to hypothesis III (*23*).

Comparison of Phrixothrix Luciferases with Firefly Luciferase Red Mutants.

20   Other residues whose mutation in firefly luciferases results in red light (*26-28*) were compared with Ph_{RE} and Pv_{GR} luciferases. Ser286, whose substitution by Asn in the firefly luciferase results in a red shift (*26*), was found to be conserved in Ph_{RE} luciferase, but replaced by Thr in Pv_{GR} and *R. ohbai* luciferases. Because Ser and Thr have very similar properties, these distinct residues do not seem to be related with

25 determination of bioluminescence color in *Phrixothrix* luciferases. Gly 326, which is replaced by Ser in the firefly red mutant CM-2 (*26*), was found to be conserved among firefly luciferases, but it was replaced by Ala in all railroad-worm and click-beetle luciferases studied. Thus these substitutions cannot account for red bioluminescence in

Ph_{RE} such as in the case of firefly luciferase mutants, probably because changes at these positions affect differently the tridimensional structure of these luciferases. Indeed, the substitutions in the red mutants always involved replacements of charged or polar groups, which can potentially interact with other residues in the tertiary structure of the

5   firefly luciferase through electrostatic interactions or by hydrogen bonding, to keep a conformation necessary for yellow-green light emission. Also, no particular similarities were found between Ph_{RE} luciferase and *P. plagiophthalamus* orange light-emitting isoenzyme, which emits closer bioluminescence color (λ_{max} = 593 nm). Ph_{RE} and Pv_{GR} luciferases showed independently the same degree of homology with both *P.*

10   *plagiophthalamus* green and orange bioluminescence-emitting isoenzymes.

Nucleic Acids of the Invention.

In view of the foregoing, the invention pertains to isolated nucleic acid molecules comprising nucleotide sequences encoding railroad worm luciferases, and the complement of these isolated nucleic acid molecules. The nucleic acid molecules can

15   be double-stranded or single-stranded; single stranded nucleic acid molecules can be either the coding (sense) strand or the non-coding (antisense) strand. The nucleic acid molecules can additionally contain a marker sequence, for example, a nucleotide sequence which encodes a polypeptide, to assist in isolation or purification of the polypeptide. Such sequences include, but are not limited to, those which encode a

20   glutathione-S-transferase (GST) fusion protein and those which encode a hemaglutinin A (HA) peptide marker from influenza. In a preferred embodiment, the nucleic acid molecule has the sequence shown in Figs. 1 and 2 (SEQ ID NOs:1 and 3).

As used herein, an "isolated" or "substantially pure" gene or nucleic acid molecule is intended to mean a gene which is not flanked by nucleotide sequences

25   which normally (in nature) flank the gene (as in other genomic sequences). Thus, an isolated gene can include a gene which is synthesized chemically or by recombinant means. Thus, recombinant DNA contained in a vector are included in the definition of "isolated" as used herein. Also, isolated nucleotide sequences include recombinant

DNA molecules in heterologous host cells, as well as partially or substantially purified DNA molecules in solution. Such isolated nucleotide sequences are useful in the manufacture of the encoded protein, as probes for isolating homologous sequences (e.g., from other mammalian species), for gene mapping (e.g., by *in situ* hybridization with

5   chromosomes), or for detecting expression of the luciferase gene in tissue (e.g., human tissue), such as by Northern blot analysis.

The present invention also encompasses variations of the nucleic acid sequences of the invention. Such variations can be naturally-occurring, such as in the case of allelic variation, or non-naturally-occurring, such as those induced by various mutagens

10   and mutagenic processes. Intended variations include, but are not limited to, addition, deletion and substitution of one or more nucleotides which can result in conservative or non-conservative amino acid changes, including additions and deletions. Preferably, the nucleotide or amino acid variations are silent or conserved; that is, they do not alter the characteristics or activity of the railroad-worm luciferases described herein.

15   Other alterations of the nucleic acid molecules of the invention can include, for example, labeling, methylation, internucleotide modifications such as uncharged linkages (e.g., methyl phosphonates, phosphotriesters, phosphoamidates, carbamates), charged linkages (e.g., phosphorothioates, phosphorodithioates), pendent moieties (e.g., polypeptides), intercalators (e.g., acridine, psoralen), chelators, alkylators, and modified

20   linkages (e.g., alpha anomeric nucleic acids). Also included are synthetic molecules that mimic nucleic acid molecules in the ability to bind to a designated sequences via hydrogen bonding and other chemical interactions. Such molecules include, for example, those in which peptide linkages substitute for phosphate linkages in the backbone of the molecule.

25   The invention also relates to fragments of the isolated nucleic acid molecules described herein. The term "fragment" is intended to encompass a portion of a nucleic acid sequence described herein which is from at least about 25 contiguous nucleotides to at least about 50 contiguous nucleotides or longer in length. One or more introns can also be present. Such fragments are useful as probes, e.g., for diagnostic methods, as

described below and also as primers or probes. Particularly preferred primers and probes selectively hybridize to a nucleic acid molecule containing the luciferase genes described herein.

5 The invention also pertains to nucleic acid molecules which hybridize under high stringency hybridization conditions, such as for selective hybridization, to a nucleotide sequence described herein (e.g., nucleic acid molecules which specifically hybridize to a nucleic acid containing the luciferase genes described herein). Hybridization probes are oligonucleotides which bind in a base-specific manner to a complementary strand of nucleic acid. Suitable probes include polypeptide nucleic

10 acids, as described in (Nielsen *et al.*, *Science 254*:1497-1500 (1991)).

Such nucleic acid molecules can be detected and/or isolated by specific hybridization (e.g., under high stringency conditions). "Stringency conditions" for hybridization is a term of art which refers to the incubation and wash conditions, e.g., conditions of temperature and buffer concentration, which permit hybridization of a

15 particular nucleic acid to a second nucleic acid; the first nucleic acid may be perfectly (i.e., 100%) complementary to the second, or the first and second may share some degree of complementarity which is less than perfect (e.g., 60%, 75%, 85%, 95%). For example, certain high stringency conditions can be used which distinguish perfectly complementary nucleic acids from those of less complementarity.

20 "High stringency conditions", "moderate stringency conditions" and "low stringency conditions" for nucleic acid hybridizations are explained on pages 2.10.1-2.10.16 and pages 6.3.1-6 in *Current Protocols in Molecular Biology* (Ausubel, F.M. *et al.*, *"Current Protocols in Molecular Biology"*, John Wiley & Sons, (1998)) the teachings of which are hereby incorporated by reference. The exact conditions which

25 determine the stringency of hybridization depend not only on ionic strength (e.g., 0.2XSSC, 0.1XSSC), temperature (e.g., room temperature, 42°C, 68°C) and the concentration of destabilizing agents such as formamide or denaturing agents such as SDS, but also on factors such as the length of the nucleic acid sequence, base composition, percent mismatch between hybridizing sequences and the frequency of

occurrence of subsets of that sequence within other non-identical sequences. Thus, high, moderate or low stringency conditions can be determined empirically. By varying hybridization conditions from a level of stringency at which no hybridization occurs to a level at which hybridization is first observed, conditions which will allow a given

5   sequence to hybridize (e.g., selectively) with the most similar sequences in the sample can be determined.

Exemplary conditions are described in Krause, M.H. and S.A. Aaronson, *Methods in Enzymology, 200*:546-556 (1991). Also, in, Ausubel, *et al.*, *"Current Protocols in Molecular Biology"*, John Wiley & Sons, (1998), which describes the

10   determination of washing conditions for moderate or low stringency conditions. Washing is the step in which conditions are usually set so as to determine a minimum level of complementarity of the hybrids. Generally, starting from the lowest temperature at which only homologous hybridization occurs, each °C by which the final wash temperature is reduced (holding SSC concentration constant) allows an increase

15   by 1% in the maximum extent of mismatching among the sequences that hybridize. Generally, doubling the concentration of SSC results in an increase in T_m of ~17°C. Using these guidelines, the washing temperature can be determined empirically for high, moderate or low stringency, depending on the level of mismatch sought.

For example, a low stringency wash can comprise washing in a solution

20   containing 0.2XSSC/0.1% SDS for 10 min at room temperature; a moderate stringency wash can comprise washing in a prewarmed solution (42°C) solution containing 0.2XSSC/0.1% SDS for 15 min at 42°C; and a high stringency wash can comprise washing in prewarmed (68°C) solution containing 0.1XSSC/0.1%SDS for 15 min at 68°C. Furthermore, washes can be performed repeatedly or sequentially to obtain a

25   desired result as known in the art. Equivalent conditions can be determined by varying one or more of the parameters given as an example, as known in the art, while maintaining a similar degree of identity or similarity between the target nucleic acid molecule and the primer or probe used.

Hybridizable nucleic acid molecules are useful as probes and primers, e.g., for diagnostic applications, as described below. As used herein, the term "primer" refers to a single-stranded oligonucleotide which acts as a point of initiation of template-directed DNA synthesis under appropriate conditions (*e.g.*, in the presence of four different nucleoside triphosphates and an agent for polymerization, such as, DNA or RNA polymerase or reverse transcriptase) in an appropriate buffer and at a suitable temperature. The appropriate length of a primer depends on the intended use of the primer, but typically ranges from 15 to 30 nucleotides. Short primer molecules generally require cooler temperatures to form sufficiently stable hybrid complexes with the template. A primer need not reflect the exact sequence of the template, but must be sufficiently complementary to hybridize with a template. The term "primer site" refers to the area of the target DNA to which a primer hybridizes. The term "primer pair" refers to a set of primers including a 5' (upstream) primer that hybridizes with the 5' end of the DNA sequence to be amplified and a 3' (downstream) primer that hybridizes with the complement of the 3' end of the sequence to be amplified.

The invention also pertains to nucleotide sequences which have a substantial identity with the nucleotide sequences described herein; particularly preferred are nucleotide sequences which have at least about 70%, and more preferably at least about 80% identity, and even more preferably at least about 90% identity, with nucleotide sequences described herein. Particularly preferred in this instance are nucleotide sequences encoding railroad-worm luciferases.

To determine the percent identity of two nucleotide sequences, the sequences are aligned for optimal comparison purposes (e.g., gaps can be introduced in the sequence of a first nucleotide sequence). The nucleotides at corresponding nucleotide positions are then compared. When a position in the first sequence is occupied by the same nucleotide as the corresponding position in the second sequence, then the molecules are identical at that position. The percent identity between the two sequences is a function of the number of identical positions shared by the sequences (i.e., % identity = # of identical positions/total # of positions x 100).

The determination of percent identity between two sequences can be accomplished using a mathematical algorithm. A preferred, non-limiting example of a mathematical algorithm utilized for the comparison of two sequences is the algorithm of Karlin *et al.* (*Proc. Natl. Acad. Sci. USA*, *90*:5873-5877 (1993)). Such an algorithm is

5   incorporated into the NBLAST program which can be used to identify sequences having the desired identity to nucleotide sequences of the invention. To obtain gapped alignments for comparison purposes, Gapped BLAST can be utilized as described in Altschul *et al.* (*Nucleic Acids Res, 25*:3389-3402 (1997)). When utilizing BLAST and Gapped BLAST programs, the default parameters of the respective programs (e.g.,

10   NBLAST) can be used. See http://www.ncbi.nlm.nih.gov. In one embodiment, parameters for sequence comparison can be set at $W=12$. Parameters can also be varied (e.g., $W=5$ or $W=20$). The value "W" determines how many continuous nucleotides must be identical for the program to identify two sequences as containing regions of identity.

15   In a related aspect, the nucleic acid fragments of the invention are used as probes or primers in assays such as those described herein. "Probes" are oligonucleotides that hybridize in a base-specific manner to a complementary strand of nucleic acid molecules. Such probes include polypeptide nucleic acids, as described in Nielsen *et al.*, *Science, 254*, 1497-1500 (1991). Typically, a probe comprises a region of

20   nucleotide sequence that hybridizes under highly stringent conditions to at least about 15, typically about 20-25, and more typically about 40, 50 or 75, consecutive nucleotides of a nucleic acid molecule comprising a nucleotide sequence selected from SEQ ID NOs: 1 or 3, the complement of SEQ ID NOs: 1 or 3. More typically, the probe further comprises a label, *e.g.*, radioisotope, fluorescent compound, enzyme, or enzyme

25   co-factor.

The invention also provides expression vectors containing a nucleic acid comprising the luciferase genes, operatively linked to at least one regulatory sequence. Many such vectors are commercially available, and other suitable vectors can be readily prepared by the skilled artisan. "Operatively linked" is intended to mean that the nucleic

acid sequence is linked to a regulatory sequence in a manner which allows expression of the nucleic acid sequence. Regulatory sequences are art-recognized and are selected to produce the luciferases described herein. Accordingly, the term "regulatory sequence" includes promoters, enhancers, and other expression control elements such as those

5   described in Goeddel, *Gene Expression Technology: Methods in Enzymology 185*, Academic Press, San Diego, CA (1990). For example, the native regulatory sequences or regulatory sequences native to the transformed host cell can be employed. It should be understood that the design of the expression vector may depend on such factors as the choice of the host cell to be transformed and/or the receptor desired to be expressed.

10   For instance, the gene of the present invention can be expressed by ligating the gene into a vector suitable for expression in either prokaryotic cells, eukaryotic cells or both (see, for example, Broach, *et al.*, *Experimental Manipulation of Gene Expression*, ed. M. Inouye (Academic Press, 1983) p. 83; *Molecular Cloning: A Laboratory Manual*, 2nd Ed., ed. Sambrook *et al.* (Cold Spring Harbor Laboratory Press, 1989) Chapters 16 and

15   17). Typically, expression constructs will contain one or more selectable markers, including, but not limited to, the gene that encodes dihydrofolate reductase and the genes that confer resistance to neomycin, tetracycline, ampicillin, chloramphenicol, kanamycin and streptomycin resistance. Vectors can also include, for example, an autonomously replicating sequence (ARS), expression control sequences, ribosome-

20   binding sites, RNA splice sites, polyadenylation sites, transcriptional terminator sequences, secretion signals and mRNA stabilizing sequences.

Prokaryotic and eukaryotic host cells transformed by the described vectors are also provided by this invention. For instance, cells which can be transformed with the vectors of the present invention include, but are not limited to, bacterial cells such as *E.*

25   *coli* (e.g., *E. coli* K12 strains), *Streptomyces*, *Pseudomonas*, *Serratia marcescens* and *Salmonella typhimurium*, insect cells (baculovirus), including *Drosophila*, fungal cells, such as yeast cells, plant cells and mammalian cells, such as thymocytes, Chinese hamster ovary cells (CHO), and COS cells. The host cells can be transformed by the described vectors by various methods (e.g., electroporation, transfection using calcium

chloride, rubidium chloride, calcium phosphate, DEAE-dextran, or other substances; microprojectile bombardment; lipofection, infection where the vector is an infectious agent such as a retroviral genome, and other methods), depending on the type of cellular host.

5   The nucleic acid molecules of the present invention can be produced, for example, by replication in a suitable host cell, as described above. Alternatively, the nucleic acid molecules can also be produced by chemical synthesis.

The nucleotide sequences of the nucleic acid molecules described herein (e.g., a nucleic acid molecule comprising SEQ ID NO:1 or SEQ ID NO:3) can be amplified by

10   methods known in the art. For example, this can be accomplished by e.g., PCR. *See generally PCR Technology: Principles and Applications for DNA Amplification* (ed. H.A. Erlich, Freeman Press, NY, NY, 1992); *PCR Protocols: A Guide to Methods and Applications* (eds. Innis, *et al.*, Academic Press, San Diego, CA, 1990); Mattila *et al.*, *Nucleic Acids Res.* 19, 4967 (1991); Eckert *et al.*, *PCR Methods and Applications* 1, 17

15   (1991); *PCR* (eds. McPherson *et al.*, IRL Press, Oxford); and U.S. Patent 4,683,202.

Other suitable amplification methods include the ligase chain reaction (LCR) (see Wu and Wallace, *Genomics 4*:560 (1989), Landegren *et al.*, *Science 241*:1077 (1988), transcription amplification (Kwoh *et al.*, *Proc. Natl. Acad. Sci. USA 86*:1173 (1989)), and self-sustained sequence replication (Guatelli *et al.*, *Proc. Nat. Acad. Sci.*

20   *USA, 87*:1874 (1990)) and nucleic acid based sequence amplification (NASBA). The latter two amplification methods involve isothermal reactions based on isothermal transcription, which produce both single stranded RNA (ssRNA) and double stranded DNA (dsDNA) as the amplification products in a ratio of about 30 or 100 to 1, respectively.

25   The amplified DNA can be radiolabeled and used as a probe for screening a library or other suitable vector to identify homologous nucleotide sequences. Corresponding clones can be isolated, DNA can be obtained following *in vivo* excision, and the cloned insert can be sequenced in either or both orientations by art recognized methods, to identify the correct reading frame encoding a protein of the appropriate

molecular weight. For example, the direct analysis of the nucleotide sequence of homologous nucleic acid molecules of the present invention can be accomplished using either the dideoxy chain termination method or the Maxam Gilbert method (see Sambrook *et al.*, *Molecular Cloning, A Laboratory Manual* (2nd Ed., CSHP, New York

5   1989); Zyskind *et al.*, *Recombinant DNA Laboratory Manual*, (Acad. Press, 1988)). Using these or similar methods, the protein(s) and the DNA encoding the protein can be isolated, sequenced and further characterized.

Antisense nucleic acid molecules of the invention can be designed using the nucleotide sequences of the present invention and the complements thereof, and

10   constructed using chemical synthesis and enzymatic ligation reactions using procedures known in the art. For example, an antisense nucleic acid molecule (*e.g.*, an antisense oligonucleotide) can be chemically synthesized using naturally occurring nucleotides or variously modified nucleotides designed to increase the biological stability of the molecules or to increase the physical stability of the duplex formed between the

15   antisense and sense nucleic acids, *e.g.*, phosphorothioate derivatives and acridine substituted nucleotides can be used. Alternatively, the antisense nucleic acid molecule can be produced biologically using an expression vector into which a nucleic acid molecule has been subcloned in an antisense orientation (*i.e.*, RNA transcribed from the inserted nucleic acid molecule will be of an antisense orientation to a target nucleic acid

20   of interest).

The host cells of the invention can also be used to produce nonhuman transgenic animals. For example, in one embodiment, a host cell of the invention is a fertilized oocyte or an embryonic stem cell into which a nucleic acid molecule of the invention has been introduced. Such host cells can then be used to create non-human transgenic

25   animals in which exogenous nucleotide sequences have been introduced into the genome or homologous recombinant animals in which endogenous nucleotide sequences have been altered. Such animals are useful for studying the function and/or activity of the nucleotide sequence and polypeptide encoded by the sequence and for identifying and/or evaluating modulators of their activity. As used herein, a "transgenic

animal" is a non-human animal, preferably a mammal, more preferably a rodent such as a rat or mouse, in which one or more of the cells of the animal includes a transgene. Other examples of transgenic animals include non-human primates, sheep, dogs, cows, goats, chickens and amphibians. A transgene is exogenous DNA which is integrated into the genome of a cell from which a transgenic animal develops and which remains in the genome of the mature animal, thereby directing the expression of an encoded gene product in one or more cell types or tissues of the transgenic animal. As used herein, an "homologous recombinant animal" is a non-human animal, preferably a mammal, more preferably a mouse, in which an endogenous gene has been altered by homologous recombination between the endogenous gene and an exogenous DNA molecule introduced into a cell of the animal, e.g., an embryonic cell of the animal, prior to development of the animal.

Methods for generating transgenic animals via embryo manipulation and microinjection, particularly animals such as mice, have become conventional in the art and are described, for example, in U.S. Patent Nos. 4,736,866 and 4,870,009, U.S. Patent No. 4,873,191 and in Hogan, *Manipulating the Mouse Embryo* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1986). Methods for constructing homologous recombination vectors and homologous recombinant animals are described further in Bradley (1991) *Current Opinion in Bio/Technology, 2*:823-829 and in PCT Publication Nos. WO 90/11354, WO 91/01140, WO 92/0968, and WO 93/04169. Clones of the non-human transgenic animals described herein can also be produced according to the methods described in Wilmut *et al.* (1997) *Nature, 385*:810-813 and PCT Publication Nos. WO 97/07668 and WO 97/07669.

Peptides and Proteins of the Invention

The present invention also provides isolated polypeptides and variants and fragments thereof that are encoded by the nucleic acid molecules of the invention. For example, as described above, the nucleotide sequences can be used to design primers to clone and express cDNAs encoding the polypeptides of the invention.

As used herein, a polypeptide is said to be "isolated" or "purified" when it is substantially free of cellular material when it is isolated from recombinant and non-recombinant cells, or free of chemical precursors or other chemicals when it is chemically synthesized. A polypeptide, however, can be joined to another polypeptide

5   with which it is not normally associated in a cell and still be "isolated" or "purified."

The polypeptides of the invention can be purified to homogeneity. It is understood, however, that preparations in which the polypeptide is not purified to homogeneity are useful. The critical feature is that the preparation allows for the desired function of the polypeptide, even in the presence of considerable amounts of

10   other components. Thus, the invention encompasses various degrees of purity. In one embodiment, the language "substantially free of cellular material" includes preparations of the polypeptide having less than about 30% (by dry weight) other proteins (*i.e.*, contaminating protein), less than about 20% other proteins, less than about 10% other proteins, or less than about 5% other proteins.

15   When a polypeptide is recombinantly produced, it can also be substantially free of culture medium, *i.e.*, culture medium represents less than about 20%, less than about 10%, or less than about 5% of the volume of the protein preparation. The language "substantially free of chemical precursors or other chemicals" includes preparations of the polypeptide in which it is separated from chemical precursors or other chemicals

20   that are involved in its synthesis. In one embodiment, the language "substantially free of chemical precursors or other chemicals" includes preparations of the polypeptide having less than about 30% (by dry weight) chemical precursors or other chemicals, less than about 20% chemical precursors or other chemicals, less than about 10% chemical precursors or other chemicals, or less than about 5% chemical precursors or other

25   chemicals.

In one embodiment, a polypeptide comprises an amino acid sequence encoded by a nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 3 and complements and portions thereof. However, the invention also encompasses sequence variants. Variants include a

substantially homologous protein encoded by the same genetic locus in an organism, *i.e.*, an allelic variant. Variants also encompass proteins derived from other genetic loci in an organism, but having substantial homology to a polypeptide encoded by a nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of

5   SEQ ID NO: 1, SEQ ID NO: 3 and complements and portions thereof. Variants also include proteins substantially homologous or identical to these polypeptides but derived from another organism, *i.e.*, an ortholog. Variants also include proteins that are substantially homologous or identical to these polypeptides that are produced by chemical synthesis. Variants also include proteins that are substantially homologous or

10   identical to these polypeptides that are produced by recombinant methods.

As used herein, two proteins (or a region of the proteins) are substantially homologous or identical when the amino acid sequences are at least about 45-55%, typically at least about 70-75%, more typically at least about 80-85%, and most typically greater than about 93% or more homologous or identical. A substantially homologous

15   amino acid sequence, according to the present invention, will be encoded by a nucleic acid molecule hybridizing to SEQ ID NO: 1, SEQ ID NO: 3 or portion thereof, under stringent conditions as more particularly described above.

To determine the percent homology or identity of two amino acid sequences, or of two nucleic acid sequences, the sequences are aligned for optimal comparison

20   purposes (*e.g.*, gaps can be introduced in the sequence of one protein or nucleic acid molecule for optimal alignment with the other protein or nucleic acid molecule). The amino acid residues or nucleotides at corresponding amino acid positions or nucleotide positions are then compared. When a position in one sequence is occupied by the same amino acid residue or nucleotide as the corresponding position in the other sequence,

25   then the molecules are homologous at that position. As used herein, amino acid or nucleic acid "homology" is equivalent to amino acid or nucleic acid "identity". The percent homology between the two sequences is a function of the number of identical positions shared by the sequences (*i.e.*, percent homology equals the number of identical positions/total number of positions times 100).

The invention also encompasses polypeptides having a lower degree of identity but having sufficient similarity so as to perform one or more of the same functions performed by a polypeptide encoded by a nucleic acid molecule of the invention. Similarity is determined by conserved amino acid substitution. Such substitutions are those that substitute a given amino acid in a polypeptide by another amino acid of like characteristics. Conservative substitutions are likely to be phenotypically silent. Typically seen as conservative substitutions are the replacements, one for another, among the aliphatic amino acids Ala, Val, Leu and Ile; interchange of the hydroxyl residues Ser and Thr, exchange of the acidic residues Asp and Glu, substitution between the amide residues Asn and Gln, exchange of the basic residues Lys and Arg and replacements among the aromatic residues Phe and Tyr. Guidance concerning which amino acid changes are likely to be phenotypically silent are found in Bowie *et al.*, Science 247:1306-1310 (1990).

A variant polypeptide can differ in amino acid sequence by one or more substitutions, deletions, insertions, inversions, fusions, and truncations or a combination of any of these. Further, variant polypeptides can be fully functional or can lack function in one or more activities. Fully functional variants typically contain only conservative variation or variation in non-critical residues or in non-critical regions. Functional variants can also contain substitution of similar amino acids that result in no change or an insignificant change in function. Alternatively, such substitutions may positively or negatively affect function to some degree. Non-functional variants typically contain one or more non-conservative amino acid substitutions, deletions, insertions, inversions, or truncation or a substitution, insertion, inversion, or deletion in a critical residue or critical region.

Amino acids that are essential for function can be identified by methods known in the art, such as site-directed mutagenesis or alanine-scanning mutagenesis (Cunningham *et al.*, *Science, 244*:1081-1085 (1989)). The latter procedure introduces single alanine mutations at every residue in the molecule. The resulting mutant molecules are then tested for biological activity *in vitro*, or *in vitro* proliferative activity.

Sites that are critical for polypeptide activity can also be determined by structural analysis such as crystallization, nuclear magnetic resonance or photoaffinity labeling (Smith *et al.*, *J. Mol. Biol.*, *224*:899-904 (1992); de Vos *et al. Science*, *255*:306-312 (1992)).

5   The invention also includes polypeptide fragments of the polypeptides of the invention. Fragments can be derived from a polypeptide encoded by a nucleic acid molecule comprising SEQ ID NO: 1, SEQ ID NO: 3 or a portion thereof and the complements thereof. However, the invention also encompasses fragments of the variants of the polypeptides described herein. As used herein, a fragment comprises at

10   least 6 contiguous amino acids. Useful fragments include those that retain one or more of the biological activities of the polypeptide as well as fragments that can be used as an immunogen to generate polypeptide-specific antibodies.

      Biologically active fragments (peptides which are, for example, 6, 9, 12, 15, 16, 20, 30, 35, 36, 37, 38, 39, 40, 50, 100 or more amino acids in length) can comprise a

15   domain, segment, or motif that has been identified by analysis of the polypeptide sequence using well-known methods, *e.g.*, signal peptides, extracellular domains, one or more transmembrane segments or loops, ligand binding regions, zinc finger domains, DNA binding domains, acylation sites, glycosylation sites, or phosphorylation sites.

      Fragments can be discrete (not fused to other amino acids or polypeptides) or

20   can be within a larger polypeptide. Further, several fragments can be comprised within a single larger polypeptide. In one embodiment a fragment designed for expression in a host can have heterologous pre- and pro-polypeptide regions fused to the amino terminus of the polypeptide fragment and an additional region fused to the carboxyl terminus of the fragment.

25   The invention thus provides chimeric or fusion proteins. These comprise a polypeptide of the invention operatively linked to a heterologous protein having an amino acid sequence not substantially homologous to the polypeptide. "Operatively linked" indicates that the polypeptide protein and the heterologous protein are fused in-frame. The heterologous protein can be fused to the N-terminus or C-terminus of the

polypeptide. In one embodiment the fusion protein does not affect function of the polypeptide *per se*. For example, the fusion protein can be a GST-fusion protein in which the polypeptide sequences are fused to the C-terminus of the GST sequences. Other types of fusion proteins include, but are not limited to, enzymatic fusion proteins,

5 for example beta-galactosidase fusions, yeast two-hybrid GAL fusions, poly-His fusions and Ig fusions. Such fusion proteins, particularly poly-His fusions, can facilitate the purification of recombinant polypeptide. In certain host cells (*e.g.*, mammalian host cells), expression and/or secretion of a protein can be increased by using a heterologous signal sequence. Therefore, in another embodiment, the fusion protein contains a

10 heterologous signal sequence at its N-terminus.

EP-A-O 464 533 discloses fusion proteins comprising various portions of immunoglobulin constant regions. The Fc is useful in therapy and diagnosis and thus results, for example, in improved pharmacokinetic properties (EP-A 0232 262). In drug discovery, for example, human proteins have been fused with Fc portions for the

15 purpose of high-throughput screening assays to identify antagonists. Bennett *et al.*, *Journal of Molecular Recognition, 8*:52-58 (1995) and Johanson *et al.*, *The Journal of Biological Chemistry, 270*,16:9459-9471 (1995). Thus, this invention also encompasses soluble fusion proteins containing a polypeptide of the invention and various portions of the constant regions of heavy or light chains of immunoglobulins of various subclass

20 (IgG, IgM, IgA, IgE).

A chimeric or fusion protein can be produced by standard recombinant DNA techniques. For example, DNA fragments coding for the different protein sequences are ligated together in-frame in accordance with conventional techniques. In another embodiment, the fusion gene can be synthesized by conventional techniques including

25 automated DNA synthesizers. Alternatively, PCR amplification of nucleic acid fragments can be carried out using anchor primers which give rise to complementary overhangs between two consecutive nucleic acid fragments which can subsequently be annealed and re-amplified to generate a chimeric nucleic acid sequence (see Ausubel *et al.*, *Current Protocols in Molecular Biology*, 1992). Moreover, many expression

vectors are commercially available that already encode a fusion moiety (*e.g.*, a GST protein). A nucleic acid molecule encoding a polypeptide of the invention can be cloned into such an expression vector such that the fusion moiety is linked in-frame to the polypeptide protein.

5   The isolated polypeptide can be purified from cells that naturally express it, purified from cells that have been altered to express it (recombinant), or synthesized using known protein synthesis methods. In one embodiment, the protein is produced by recombinant DNA techniques. For example, a nucleic acid molecule encoding the polypeptide is cloned into an expression vector, the expression vector introduced into a

10   host cell and the protein expressed in the host cell. The protein can then be isolated from the cells by an appropriate purification scheme using standard protein purification techniques.

      In general, polypeptides or proteins of the present invention can be used as a molecular weight marker on SDS-PAGE gels or on molecular sieve gel filtration

15   columns using art-recognized methods. The polypeptides of the present invention can be used to raise antibodies or to elicit an immune response. The polypeptides can also be used as a reagent, *e.g.*, a labeled reagent, in assays to quantitatively determine levels of the protein or a molecule to which it binds (*e.g.*, a receptor or a ligand) in biological fluids. The polypeptides can also be used as markers for cells or tissues in which the

20   corresponding protein is preferentially expressed, either constitutively, during tissue differentiation, or in a diseased state. The polypeptides can be used to isolate a corresponding binding partner, *e.g.*, receptor or ligand, such as, for example, in an interaction trap assay, and to screen for peptide or small molecule antagonists or agonists of the binding interaction.

25   In another aspect, the invention provides antibodies to the polypeptides and polypeptide fragments of the invention, *e.g.*, having an amino acid sequence encoded by a nucleic acid molecule comprising all or a portion of SEQ ID NO: 1 or SEQ ID NO: 3. The term "antibody" as used herein refers to immunoglobulin molecules and immunologically active portions of immunoglobulin molecules, *i.e.*, molecules that

contain an antigen binding site that specifically binds an antigen. A molecule that specifically binds to a polypeptide of the invention is a molecule that binds to that polypeptide or a fragment thereof, but does not substantially bind other molecules in a sample, *e.g.*, a biological sample, which naturally contains the polypeptide. Examples

5 of immunologically active portions of immunoglobulin molecules include F(ab) and F(ab')$_2$ fragments which can be generated by treating the antibody with an enzyme such as pepsin. The invention provides polyclonal and monoclonal antibodies that bind to a polypeptide of the invention. The term "monoclonal antibody" or "monoclonal antibody composition", as used herein, refers to a population of antibody molecules that contain

10 only one species of an antigen binding site capable of immunoreacting with a particular epitope of a polypeptide of the invention. A monoclonal antibody composition thus typically displays a single binding affinity for a particular polypeptide of the invention with which it immunoreacts.

Polyclonal antibodies can be prepared as described above by immunizing a

15 suitable subject with a desired immunogen, *e.g.*, polypeptide of the invention or fragment thereof. The antibody titer in the immunized subject can be monitored over time by standard techniques, such as with an enzyme linked immunosorbent assay (ELISA) using immobilized polypeptide. If desired, the antibody molecules directed against the polypeptide can be isolated from the mammal (*e.g.*, from the blood) and

20 further purified by well-known techniques, such as protein A chromatography to obtain the IgG fraction. At an appropriate time after immunization, *e.g.*, when the antibody titers are highest, antibody-producing cells can be obtained from the subject and used to prepare monoclonal antibodies by standard techniques, such as the hybridoma technique originally described by Kohler and Milstein (1975) *Nature, 256*:495-497, the human B

25 cell hybridoma technique (Kozbor *et al.* (1983) *Immunol. Today, 4*:72), the EBV-hybridoma technique (Cole *et al.* (1985), *Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, Inc., pp. 77-96) or trioma techniques. The technology for producing hybridomas is well known (see generally *Current Protocols in Immunology* (1994) Coligan *et al.* (eds.) John Wiley & Sons, Inc., New York, NY). Briefly, an

immortal cell line (typically a myeloma) is fused to lymphocytes (typically splenocytes) from a mammal immunized with an immunogen as described above, and the culture supernatants of the resulting hybridoma cells are screened to identify a hybridoma producing a monoclonal antibody that binds a polypeptide of the invention.

5   Any of the many well known protocols used for fusing lymphocytes and immortalized cell lines can be applied for the purpose of generating a monoclonal antibody to a polypeptide of the invention (see, *e.g.*, *Current Protocols in Immunology, supra*; Galfre *et al.* (1977) *Nature, 266*:55052; R.H. Kenneth, in *Monoclonal Antibodies: A New Dimension In Biological Analyses*, Plenum Publishing Corp., New

10   York, New York (1980); and Lerner (1981) *Yale J. Biol. Med., 54*:387-402. Moreover, the ordinarily skilled worker will appreciate that there are many variations of such methods that also would be useful.

Alternative to preparing monoclonal antibody-secreting hybridomas, a monoclonal antibody to a polypeptide of the invention can be identified and isolated by

15   screening a recombinant combinatorial immunoglobulin library (*e.g.*, an antibody phage display library) with the polypeptide to thereby isolate immunoglobulin library members that bind the polypeptide. Kits for generating and screening phage display libraries are commercially available (*e.g.*, the Pharmacia *Recombinant Phage Antibody System*, Catalog No. 27-9400-01; and the Stratagene *Surf*ZAP™ Phage Display Kit, Catalog No.

20   240612). Additionally, examples of methods and reagents particularly amenable for use in generating and screening antibody display library can be found in, for example, U.S. Patent No. 5,223,409; PCT Publication No. WO 92/18619; PCT Publication No. WO 91/17271; PCT Publication No. WO 92/20791; PCT Publication No. WO 92/15679; PCT Publication No. WO 93/01288; PCT Publication No. WO 92/01047; PCT

25   Publication No. WO 92/09690; PCT Publication No. WO 90/02809; Fuchs *et al.* (1991) *Bio/Technology, 9*:1370-1372; Hay *et al.* (1992) *Hum. Antibod. Hybridomas, 3*:81-85; Huse *et al.* (1989) *Science, 246*:1275-1281; Griffiths *et al.* (1993) *EMBO J., 12*:725-734.

Additionally, recombinant antibodies, such as chimeric and humanized monoclonal antibodies, comprising both human and non-human portions, which can be made using standard recombinant DNA techniques, are within the scope of the invention. Such chimeric and humanized monoclonal antibodies can be produced by

5   recombinant DNA techniques known in the art.

In general, antibodies of the invention (*e.g.*, a monoclonal antibody) can be used to isolate a polypeptide of the invention by standard techniques, such as affinity chromatography or immunoprecipitation. A polypeptide-specific antibody can facilitate the purification of natural polypeptide from cells and of recombinantly produced

10   polypeptide expressed in host cells. Moreover, an antibody specific for a polypeptide of the invention can be used to detect the polypeptide (*e.g.*, in a cellular lysate, cell supernatant, or tissue sample) in order to evaluate the abundance and pattern of expression of the polypeptide. Antibodies can be used diagnostically to monitor protein levels in tissue as part of a clinical testing procedure, *e.g.*, to, for example, determine the

15   efficacy of a given treatment regimen. Detection can be facilitated by coupling the antibody to a detectable substance. Examples of detectable substances include various enzymes, prosthetic groups, fluorescent materials, luminescent materials, bioluminescent materials, and radioactive materials. Examples of suitable enzymes include horseradish peroxidase, alkaline phosphatase, β-galactosidase, or

20   acetylcholinesterase; examples of suitable prosthetic group complexes include streptavidin/biotin and avidin/biotin; examples of suitable fluorescent materials include umbelliferone, fluorescein, fluorescein isothiocyanate, rhodamine, dichlorotriazinylamine fluorescein, dansyl chloride or phycoerythrin; an example of a luminescent material includes luminol; examples of bioluminescent materials include

25   luciferase, luciferin, and aequorin, and examples of suitable radioactive material include 125I, 131I, 35S or 3H.

Applicability.

Nucleic acid molecules of this invention can be used as sensitive and versatile tools to report molecular events genetically regulated inside living cells and their extracts in biological assays. (See Gould and Subramani, 1988). The luciferase genes

5   can be ligated to DNA vectors (such as plasmids and virus) downstream regulator elements called promoters, and inserted inside the cells. Under the proper conditions, the promoter will drive the transcription of the luciferase gene and the expression of active luciferases in a manner that is dependent of the degree of activation of the promoter. The expression of the luciferase is directly proportional to the amount of

10   light produced in presence of exogenously supplied D-luciferin and ATP (present in the intracellular medium) and can be quantified using commercially available photometers, such as luminometers. Based on this principle, luciferase genes can be used to monitor the transcriptional activity of promoters, to diagnostic and quantify the spreading of viral infections in animal and vegetal tissues or to diagnostic the presence of virus in

15   biological samples. In this case, the luciferase gene is ligated to a DNA vector downstream a viral promoter, which drives the transcription and expression of luciferase. This system is then introduced inside the cell or the biological sample to be tested. If the biological sample is infected by the virus from which the promoter was isolated, the molecular signals resulting from such virus will activate the promoter that

20   will start to transcribe the luciferase gene and to express luciferase, whose light will confirm the infection. In addition, the amount of light produced will be proportional to the degree of promoter activation which can be used to quantify the titer of the infection (this methodology was already successfully used to detect the presence of virus such as HIV). In brief, luciferase genes can be used as reporter genes for the following

25   purposes: (1) promoter transcriptional activity determination (e.g., single and dual promoters); (2) viral diagnostics and quantification; (3) *in vivo* expression and visualization in (through) mammalian tissues such as blood, bone tissue and muscular tissue; (3) cytotoxcicity tests; (4) cell viability determination; (5) ATP quantification in biological samples ; (6) biosensor for environmental and chemical stress agents, among

others. See Gould and Subramani, 1988; Schwartz *et al.*, 1990; Contag *et al.*, 1995; Comhaire *et al.*, 1989; Stanley, 1989; Lundin *et al.*, 1989; Tauriainen *et al.*, 1999.

In the past, only the gene for the luciferase of the North-American firefly *Photinus pyralis*, which produces yellow-green light, has been widely applied as

5   reporter gene. One disadvantage of this reporter gene is that the yellow-green light is considerably absorbed in certain biological samples with high optical density in the blue-yellow region of the spectrum such as the blood and other highly pigmented biological samples. Other genes that code for beetle luciferases that emit in the green-orange region of the spectrum were also cloned (Wood, 1995), but they were not widely

10   used and they suffer similar limitations in pigmented biological samples.

The gene of the red emitting luciferase arising from *Phrixothrix* described herein (Viviani *et al.*, 1999; incorporated herein in its entirety) offer a notable wider potential for application, mainly for purposes of detecting light in biological samples with high optical density in the blue-yellow region of the spectrum such as blood, bone tissue and

15   muscular tissue among others. In addition, the simultaneous use of the genes for the green and red emitting luciferases, with the aim to monitor two different promoters (regulator elements) at the same time, offer the possibility to use 2 reporter genes whose signals can be clearly distinguished by the color of the light, with a minimum of bioluminescence spectrum overlap (<10%) avoiding interferences. Furthermore, both

20   luciferases, arising from the same bioluminescent system, work under the same biochemical conditions (in presence of ATP, magnesium ions which are natural intracellular components) and D-Luciferin which should be supplied exogenously. Since the beetle luciferases genes presently used code for luciferases which produce light in the green-orange range of the spectrum, their simultaneous application is limited

25   by the considerable degree of emission spectra overlapping (>25%), imposing severe difficulties to discriminate one signal from another. One solution to this problem was previously patented by Promega Co. (USA) (Sherf *et al.*, 1996), which used the genes for firefly luciferase which emit yellow-green light and a celenterate (jellyfish) luciferase which emit blue light. However, since these luciferases arise from different

bioluminescent systems, they work in distinct biochemical conditions, which is not always a convinient condition.

Thus, the research for efficiency is satisfied by the use of the luciferases coded by the sequences object of this patent application, since they are able to produce the same green and red bioluminescence of *Phrixothrix*, constituting a system which works in the same biochemical conditions and whose bioluminescence spectra have not overlapping.

These genes can also be used for the industrial scale production of luciferases for analytical applications. Light emitted in bioluminescent assays using luciferases is actually detectable by a range of commercially avaliable photometric instruments such as luminometers, photon counters, fluorometers, CCD-camera systems.

The invention will be further illustrated by the following exemplification which is not intended to be limiting in any way.

EXEMPLIFICATIONS

IN VIVO EXPRESSION OF LUCIFERASE BY E. COLI

MATERIALS AND METHODS.

Reagents.

Isopropyl-β-D-thiogalactopyranoside (IPTG[1]), 5-bromo-4-chloro-3-indoyl-$\hat{a}$-galactopyranoside (X-Gal), dithiothreitol (DTT), D-luciferin (sodium salt), ampicillin, tetracyclin, and kanamycin were from Wako Pure Chemicals (Osaka, Japan); coenzyme-A (CoA) and adenosine triphosphate (ATP) from Oriental Yeast Co (Osaka, Japan); Isogen reagent, restriction enzymes, and Taq polymerase from Nippon Gene (Toyama, Japan); Oligo Tex dT30 and DNA ligation kit from Takara Shuzo (Kyoto, Japan); cDNA synthesis kits from Amersham Pharmacia Biotech (Tokyo, Japan); Gigapack III Gold packaging kit from Stratagene (La Jolla, CA); and ABI PRISM Dye terminator Cycle Sequencing kit from Perkin-Elmer (Foster City, CA).

Bacterial Strains and Media.

        Echerichia coli XL1-Blue MRF' and SOLR strains were purchased from Stratagene (La Jolla, CA). *E. coli* cells were usually grown in Luria Bertani (LB) medium (1% bacto-triptone, 0.5% yeast extract, 0.5% NaCl). Cell densities were

5   measured by absorbance at 600 nm.

Insects.

        Railroad-worms were collected at night as described (*5*). Larvae of *P. vivianii* were collected in pastures at Fazenda São Francisco near Parque Nacional das Emas (prefecture of Mineiros, Goías State). Larvae of *P. hirtus* were collected into cerradão

10   formation at Fazenda Sta Cruz (Costa Rica prefecture, Mato Grosso do Sul State) (*31*) near the former place. Living specimens were cleaned with distilled water, frozen in liquid nitrogen, and stored at -80°C.

Construction and Screening of cDNA Libraries.

        cDNA libraries were constructed using methodology similar to *Pyrearinus*

15   *termitilluminans* luciferase cloning (*16*). For *P. vivianii* larvae, total RNA was extracted from the whole bodies (without head) of 8 specimens, yielding 360 μg, using Isogen reagent according established procedures (*32*). Them RNAs were isolated using Oligo-dT latex in accordance with Kuribayashi *et al.* (*33*). cDNAs were synthesized from 4 μg of isolated mRNAs using Time Saver cDNA synthesis kit. The first strand

20   reaction was carried out in the presence of oligo-dT$_{12-18}$ primer. After the synthesis of the second cDNA strand, *EcoRI/NotI* adaptors were ligated to the blunt ended cDNA. The cDNA (about 50 ng for *P. vivianii* bodies) was ligated to 1 μg of *EcoRI* pre-digested/dephosphorilated λZAP II (Stratagene, La Jolla, CA) vector in a volume of 5 μL of ligation reaction mixture (1 mM ATP, 7 mM MgCl$_2$, 1 mM DTT in 50 mM Tris-

25   HCl, pH 8.0, and 1 Weiss unit of T4 ligase) overnight at 16°C. The ligation mixtures were then packaged using Gigapack III Gold packaging extracts. The original library of *P. vivianii* bodies (7.5 X 10^5 pfu) was then *in vivo* excised into *E. coli* XL1-Blue cells in

the presence of helper phage to obtain pBluescript (pBl) libraries. The excised

phagemids were used to transform *E. coli* SOLR cells. The plasmid library was

screened by photodetection (*34*) using a cooled-CCD camera system (ATTO; Tokyo,

Japan), after spraying 1 mM D-luciferin (0.1 M citrate buffer pH 5.0) onto IPTG induced

5   colonies at 20°C during 12 h. For *P. hirtus* head lantern library construction, total RNA

was extracted from 17 heads and 4 μg of mRNA was used to synthesize cDNA. All

other procedures were essentially similar to those described above, except the original

library (2.2 X 10^4 pfu) was further amplified (1 X 10^9 pfu) before excision of the pBl

library.

10   *Sequence Analysis.*

The cDNA for green light-emitting luciferase (Pv$_{GR}$), the 0.75 kb long

*Eco*RV/*Bam*HI and *Bam*HI/*Bam*HI fragments were subcloned into pBl and pUC

vectors, respectively (*35*). The cDNA for red light-emitting luciferase (Ph$_{RE}$) was

digested with *Eco*RI, and the 3 resulting fragments (0.8; 0.6; 0.3 kbp) were subcloned

15   into *Eco*RI digested/dephosphorilated pUC 18 vector. All of these constructions and the

original luciferase cDNA-containing plasmids were sequenced by the dydeoxy chain

termination method (*36*) using dye-labeled terminator kit specifically developed for the

ABI PRISM 377 automatic sequencer (Perkin-Elmer; Foster City-CA). For extension of

pUC vector, M$_{13}$ (-21) and reverse primers were used, whereas for pBl vector T$_7$ and

20   reverse primers were used. Three additional primers, VA$_1$ (5′-

ATGTACTTTCAATCTCTTTGCTAC-3′) (SEQ ID NO:5), VA$_3$ (5′-

AAGTCTAACTATAAGATAAGTTCTTA-3′) (SEQ ID NO:7), and VA$_4$ (5′-

CAAGTTTCAGTTAATCCATAT-3′) (SEQ ID NO:6), were designed from the known

partial sequences in order to sequence the internal regions of Ph$_{RE}$ and Pv$_{GR}$. Sequence

25   comparisons, multi-alignments, and determination of the protein hydropathy profiles,

molecular weights and isoeletric points were made using version 7.3 of Genetyx-mac

software (Software Development Co., Ltd., Tokyo, Japan).

Expression and Preparation of Luciferase Extracts.

Liquid cultures of SOLR cells carrying luciferase insert containing pBl were grown on LB/amp (50 µg/mL) medium at 37 ˚C with shaking overnight. The preculture (1/100 vol) was then inoculated in LB/amp (50 µg/mL) in the presence of 1 mM IPTG

5 and incubated at 23˚C during 24 h (OD_{600} =1.8). The cells were harvested by centrifugation at 3000 rpm during 10 min at 4˚C, and the pellet was resuspended in cold extraction buffer (0.1 M sodium phosphate buffer, pH 7.5, containing 2 mM EDTA, 1 mM DTT, and 1% Triton X-100). Lysozyme was added to the final concentration of 1 mg/mL and the suspension incubated 15 min at 0˚C and frozen at

10 -80˚C during 15 min. The lysate was centrifuged at 12000*g* at 4˚C during 15 min. The supernatant was then fractionated with ammonium sulfate. The fraction precipitated between 55% and 70% saturation was dissolved in cold extraction buffer and stored at -20˚C in the presence of glycerol 15%, to maintain the activity.

Luciferase Assays.

15 The activity levels were measured using a Luminescencer AB-2000 luminometer (Atto; Tokyo, Japan) by integration of total light emitted. The assay consisted of the addition of 50 µL of standard solution (2 mM ATP, 0.5 mM D-luciferin, 4 mM $MgSO_4$ in 0.1 M Tris-HCl buffer, pH 8.0) to 10 µL of luciferase-containing extracts at 25˚C. *In vivo* light intensities were measured after adding 50 µL of 0.5 mM D-luciferin in 0.1 M

20 sodium citrate buffer pH 5.0 to 10 µL of bacterial suspension into a microtiter plate.

Kinetic Measurements.

Measurements of light intensities for K_M determinations were made using the luminometer described above. For luciferin K_M estimation, 50 µL of 4 mM ATP solution (0.1 M Tris-HCl, pH 8.0, and 8 mM $MgSO_4$) was injected to 50 µL of crude

25 extract diluted 10 times in 0.1 M Tris-HCl buffer, pH 8.0, containing luciferin (0.03-2 mM). For ATP K_M estimation, 50 µL of 0.5 mM luciferin solution (0.1 M Tris-HCl

buffer, pH 8.0, 8 mM MgSO$_4$) was injected to 50 µL of 10 times diluted extract containing ATP (0.1-4 mM). The assays were carried out at 25°C. Each point of the Michaelis-Menten curve was assayed in quadruplicate. The K_M values were estimated by Lineweaver-Burk plots of the reciprocal of light intensities versus substrate

5 concentration.

Bioluminescence Spectra.

Emission spectra were determined using a Hitachi F4500 spectrofluorometer, supplied with a Hamamatsu Photonics R 928 F photomultiplier, with the excitation lamp shut down. The spectra were automatically corrected for the photosensitivity of

10 the equipment. For bacterial *in vivo* spectra determinations, 500 µL of bacterial suspension and 500 µL of 0.5 mM D-luciferin in 0.1 M sodium citrate buffer, pH 5.0, and 10 mM MgSO$_4$ were mixed into a cuvette in front of the emission window (*16*). *In vitro* spectra were recorded 3 min after mixing 10-100 µL of luciferase-containing extract to 900 µL of standard reaction mixture (0.5 mM D-luciferin, 2 mM ATP, 4 mM

15 MgSO$_4$, 0.5 mM CoA, and 1% Triton X-100 in 0.1 M Tris-HCl uffer, pH 8.0), to a final volume of 1 mL, into a luorometer cuvette (*16*) in front of the emission window. The pH effect on in vitro spectra was measured in 0.1 M sodium phosphate buffer (pH 6-8) instead of Tris-HCl buffer as described (*16*). The spectra measured at pH 8.0 in both buffers were essentially identical in shape.

20 RESULTS

Isolation of Positive Clones and Expression of Active Luciferases.

The *P. vivianii* body cDNA library yielded 7.5 X 10^5 recombinant plaques. After screening about 10 000 colonies, we isolated 1 positive clone for light emission (Pv$_{GR}$). The cDNA library for *P. hirtus* head lanterns yielded 2.2 X 10^4 recombinant

25 plaques and was amplified (1 X 10^9 pfu) before excision. Four positive clones were

found after screening 3000 colonies of the excised amplified library. The most intense light-emitting clone (Ph$_{RE}$) was isolated for further analysis. Upon D-luciferin spraying, the IPTG-induced SOLR colonies containing pBl-Pv$_{GR}$ and pBl-Ph$_{RE}$ displayed weak green and red bioluminescence, respectively, visible after dark-room eye adaptation.

5   The luminescence maximum intensity of pBl-Ph$_{RE}$-containing colonies was reached before and decayed sooner than that of pBl-Pv$_{GR}$-containing colonies. Luminometer measurement of the total light output of both in *in vivo* and *in vitro* bioluminescence assays of IPTG-induced colonies gave nearly the same values for Pv$_{GR}$ and Ph$_{RE}$ luciferases, after correction for the spectral photosensitivity of the equipment.

10   *cDNAs Structures and Sequences.*

    The Pv$_{GR}$ cDNA (NCBI access number: AF139644) is a 1764 bp long fragment. The start codon was found 25 bp downstream from the cDNA 5′terminus, which follows the *Pst*I restriction site of pBluescript polylinker. An open reading frame of 1635 bp, which codes for a 545 amino acid long polypeptide was found. After the

15   stop codon, a 105 bp long 3′ untranslated region followed by a terminal 26 bp poly-A tail was found.

    The Ph$_{RE}$ cDNA (NCBI access number: AF139645) is a 1760 bp long fragment. The cDNA has a 41 bp untranslated region upstream from the first ATG. The sequence of the last 10 bp before the starting codon was essentially identical to that of Pv$_{GR}$. An

20   open reading frame of 1638 bp, coding for a potential 546 residue long polypeptide was found. After the stop codon, a 61 bp long downstream untranslated region with a terminal 7 bp long poly-A tail was found.

    Protein Sequences.

    The overall identity between Pv$_{GR}$ and Ph$_{RE}$ luciferases was 71%. Pv$_{GR}$ and Ph$_{RE}$

25   luciferases showed 66.6% and 56% identity, respectively, with the Japanese railroad-worm *Ragophthalmus ohbai* luciferase, recently cloned in our laboratory (*37*). As expected, these values are very close to those observed for *Phengodes* luciferase, with

the same set of enzymes (*17*). Pv_{GR} luciferase showed a slightly higher identity with firefly luciferases (50-55%) than Ph_{RE} did with the same set of enzymes (46-49%). Both luciferases showed 47-49% identity with click-beetle luciferases. The overall identity shared with acyl CoA ligases is 25%. Like most beetle luciferases, both

5   luciferases showed the peroxissomal targeting tripeptide SKL before the stop codons (7).

Protein Properties.

The calculated molecular weights of Pv_{GR} and Ph_{RE} luciferases were 59,626 and 60,951 kDa and were close to those estimated by Western Blotting (data not shown).

10   The isoeletric points, calculated from the deduced primary structures, were 6.26 and 7.0 for Pv_{GR} and Ph_{RE} luciferases, respectively. Pv_{GR} luciferase showed a small increase in the proportion of hydrophobic and neutral residue content in relation to Ph_{RE} luciferase. The hydropathy profiles of these luciferases were similar; however there were some regions which showed major differences. In particular the region from residues 350-

15   360, which includes the additional Arg residue in Ph_{RE} luciferase, showed a major increase of hydrophobic character in Ph_{RE} luciferase in relation to Pv_{GR} luciferase.

Kinetic Parameters.

The K_M values for ammonium sulfate fractionated extracts were measured. For

20   Ph_{RE} luciferase the K_M for luciferin was 20 µM, whereas for ATP it was 240 µM. However for unknown reasons Pv_{GR} luciferase K_M values were much higher than for Ph_{RE} luciferase (150 µM for luciferin and 350 µM for ATP). The rise time to the peak of intensity and the decay rate of *in vitro* bioluminescence reaction were faster for Ph_{RE} luciferase than for Pv_{GR} (results not shown). A similar property was observed for the

25   time course of the *in vivo* bioluminescence after spraying D-luciferin.

Spectral Properties.

The *in vivo* and *in vitro* bioluminescence spectra emitted by Pv_{GR} luciferase are centered at 549 nm, thus in the green region, close to that reported for the native enzyme extracted from the larval *P. vivianii* lateral lanterns (λmax = 542 nm; Fig. 3) *(5)*, and to the spectra emitted by the recently cloned *Phengodes* (λ_{max} = 546 nm) *(17)* and *R. ohbai* (λ_{max} 556 nm) *(37)* luciferases. Both the *in vivo* and *in vitro* bioluminescence spectra of Ph_{RE} luciferase showed a peak at 622 nm (Fig. 3), thus 73 nm shifted in relation to Pv_{GR} luciferase. The in vitro spectrum is 6 nm blue-shifted in relation to that emitted by the native enzyme extracted from the laval *P. hirtus* head lanterns (λ_{max} = 628 nm) *(5)*, in part due to distinct equipment used for measurements. This spectrum is much narrower (half bandwidth = 55 nm) than that emitted by Pv_{GR} luciferase (half-bandwidth = 70.5 nm) and other bettle luciferases. The bioluminescence spectra of both recombinant luciferases did not suffer red shift upon decreasing the pH from 8.0 to 6.0, although they underwent a decrease of intensity.

While this invention has been particularly shown and described with references to preferred embodiments thereof, it will be understood by those skilled in the art that various changes in form and details may be made therein without departing from the scope of the invention encompassed by the appended claims.

REFERENCES

1. Seliger, H. H., Buck, J. B., Fastie, W. G., and McElroy, W. D., *J. Gen. Physiol.* *48*:95-104 (1964).

2. Biggley, W. H., Lloyd, J. E., and Seliger, H. H., *J. Gen. Physiol.* *50*:1681-1692 (1967).

3. Colepicolo, N. P., Costa, C., and Bechara, E. J. H., *Insect Biochem. 16*, 803-810 (1986).

4. Viviani, V. R., and Bechara, E. J. H., *Photochem. Photobiol. 58*:615-622 (1993).

5. Viviani, V. R., and Bechara, E. J. H., *Ann. Entomol. Soc. Am. 90*:389-398 (1997).

6. Lall, B. A, Seliger, H. H., Biggley, W. H., and Lloyd, J. E., *Science 210*:560-562 (1980).

7. Wood, K. V., *Photochem. Photobiol. 62*:662-673 (1995).

8. De Wet, J. R., Wood, K. V., Helinsky D. R., and DeLuca, M, Proc. Natl. Acad. Sci. U.S.A. 82:, 7870-7873 (1985).

9. Tatsumi, H., Masuda, T., Kajiyama, N., and Nakano, E., *J. Biolum. Chemilum. 3*:75-78 (1989).

10. Tatsumi, H., Kajiyama, N., and Nakano, E., Biochim. Biophys. Acta 1131:, 161-165 (1992).

11. Devine, J. H., Kutuzova, G. D., Green, V. A., Ugarova, N. N., and Baldwin, T. O., Biochim. Biophys. Acta 1173:121-132 (1993).

12. Ohmiya, Y., Ohba, N., Toh, H., and Tsuji, F. I., *Photochem. Photobiol.* 62:309-313 (1995).

13. Sala-Newby, G. B., Thomson, C. M., and Campbell, A. K. (1996) *Biochem. J. 313*, 761-767.

14. Li, Ye, Buck, L. M., Scaeffer, H. J., and Leach, F. R., *Biochim. Biophys. Acta 1339*:39-52 (1997).

15. Wood, K. V., Lam, Y. A., Seliger, H. H., and McElroy, W. D., *Science 244*: 700-702 (1989).

16. Viviani, V. R., Silva, A. C. R., Barbosa, G. N., Perez, G. L. O., Santelli, R. V., Bechara, E. J. H., and Reinach, F. C. *Proceedings of the X th International Symposium on Bioluminescence and Chemilum*inescence (04-08/09/1998) Bologna, Italy (1999).

17. Gruber, M. G., Kutuzova, G. D., and Wood, K. V., In *Bioluminescence and Chemiluminescence: Molecular Reporting with Photons.* Proceedings of the 9 th International Sym-posium (Hastings, J. W., Kricka, L. J., Stanley, P. E., Eds.) pp 244-247, John Wiley and Sons, Chichester, U.K (1996).

18.   White, E. H., Rapaport, E., Hopkins, T. A., and Seliger, H. H., *J. Am. Chem. Soc.* 91:1243-1245 (1969).

19.   White, E. H., and Branchini, B., *J. Am. Chem. Soc. 97*:2178-2180 (1975).

20.   Viviani, V. R., and Bechara, E. J. H., Photochem. Photobiol. 62:490-495 (1995).

21.   DeLuca, M. (1969) *Biochemistry 8*, 160-166.

22.   Brovko, Lyu, Dementieva, E. I., and Ugarova, N. N., In *Bioluminescence and Chemiluminescence: Molecular Re-porting with Photons.* Proceedings of the 9th International Symposium (Hastings, J. W., Kricka, L. J., Stanley, P. E., Eds.) pp 206-211, John Wiley and Sons, Chichester, U.K (1996).

23.   McCapra, F., Gilfoyle, D. J., Young, D. W., Church, N. J., and Spencer, P., In *Bioluminescence and Chemilumi-nescence: Fundamental and Applied Aspects* (Campbell, A. K., Kricka, L. J. Stanley, P. E., Eds.) pp 387-391, John Wiley and Sons, Chichester, U.K (1994).

24.   Wood, K. V., *J. Biol. Chemilum.* 5:107-114 (1990).

25.   Ohmiya, Y., Hirano, T., and Ohashi, M., *FEBS Lett. 384*:83-86 (1996).

26.   Kajiyama, N., and Nakano, E., *Protein Eng. 4*:691-693 (1991).

27.   Mamaev, S. V., Laikhter, A. L., Arslan, T., and Hecht, S. M., *J. Am. Chem. Soc. 118*:7243-7244 (1996).

28.   Ueda, H., Yamanouchi, H., Kitayama, A., Inoue, K., Hirano, T., Suzuki, E., Nagamune, T., and Ohmiya, Y., In *Bioluminescence and Chemiluminescence: Molecular Reporting with Photons.* Proceedings of the 9th International Symposium (Hastings, J. W., Kricka, L. J., Stanley, P. E., Eds.) pp 216-219, John Wiley and Sons, Chichester, U.K. (1996).

29.   Conti, E., Franks, N. P., and Brick, P., *Structure 4*:287-298 (1996).

30.   Wittmer, W., *Mitt. Entomol. Ges. BRD 42*:130-132 (1993).

31.   Costa, C., Vanin, S. A., Casari, S. A., and Viviani, V. R. *Ilheringia* (1999).

32.   Chomczynsky, P., and Sacchi, N., *Anal. Biochem. 162*:156-159 (1987).

33. Kuribayashi, K., Hirata, M., Kiraoka, O., Miyamoto, C., and Furushi, Y., *Nucleic Acids Symposium Series,* Vol. 1, p 61, Harvard Medical School, John Wiley and Sons, NewYork (1988).

34. Wood, K. V., and DeLuca, M., *Anal. Biochem. 161*:501-507 (1987).

35. Ausbel, M. F., et al., *Short Protocols in Molecular Biology.* (Ausbel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., Struhl, K., Eds.) Harvard Medical School, John Wiley and Soons, New York (1992).

36. Sambrook, J., Fritsch, E. F., and Maniatis, T., *Molecular Cloning, a Laboratory Manual,* 2nd ed. Cold Spring Harbor, New York (1989).

37. Sumiya, M., Viviani, V. R., Ohba, N., and Ohmiya, Y., *J. Biolum. Chemilum.* 13:225 (1998).

38. White, E., Rapaport, E., Seliger, H. H., and Hopkins, T. A., *Bioorg. Chem. 1*:92-122 (1971).

39. White, E., and Roswell, D. F., *Photochem. Photobiol. 53*:131-136 (1991).

40. Alter, S. C., and DeLuca, M., *Biochemistry 25*:1599-1605 (1986).

41. Branchini, B. R., Magyar, R. A., Marcantonio, K. M.,Newberry, K. J., Stroh, J. G., Hinz, L. K., and Murtiashaw, M. H., J. Biol. Chem. 272:19359-19364 (1997).

42. Branchini, B., Magyar, R. A., Murtiashaw, M. H., Anderson, S. M., and Zimmer, M., *Biochemistry 37*:15311-15319 (1998).

43. Conti, E., Stachelhaus, T., Marahiel, M. A., and Brick, P., *EMBO J.* 16:4174-4183 (1997).

44. Seliger, H. H., and McElroy, W. D., Proc. Natl. Acad. Sci. U.S.A. 52:75-81 (1964).

45. DeLuca, M., *Biochemistry 13*:921-925 (1974).

46. Bowie, L. J., Horak, V., and DeLuca, *M., Biochemistry 8*:1598-1607 (1969).

47. Campbell, A K., *Chemiluminescence: principles and applications in biology and medicine,* Ellis Horwood, Cichester, England (1988).

48.   Hopp, T. P., and. Woods, K. R., *Proc. Natl. Acad Sci. U.S.A. 78*:3824–3828 (1981).

Comhaire, F. H., L. Vermeulen, L. Monsieur and A. Hinting, "Determination of adenosine triphosphate in human semen to estimate the fertilizing potential and to
5   quantify sperm antibodies", *J. Biolum. Chemilum.* 4:399-405 (1989).

Contag, C. H., Spilman, S. D., Contag, P. R., Oshiro M., Eames B., Dennery P., Stevenson D. H. and D. A. Benaron, "Visualizing gene expression in living mammals using a bioluminescent reporter", *Photochem. Photobiol. 66*:523-531 (1997).

De Wet, J. R., K. V. Wood, D. R. Helinsky and M. DeLuca, "Cloning of firefly
10 luciferase cDNA and expression of active luciferase in *Echerichia coli.", Proc. Natl. Acad. Sci. USA 82*:7870-7873 (1985).

Gould, S. J. and S. Subramani, "Firefly luciferase as a tool in molecular and cell biology", *Anal. Biochem. 175*:5-13 (1988).

Lundin, A., H. Hallander, A. Kallner, U. K. Lundin and E. Osterberg, " Bacteriuria testing
15 by the ATP method as an integral part in the diagnosis and therapy of urinary tract infection (UTI)", *J. Biolum. Chemilum.* 4:381-389 (1989).

Schwartz O., J. L. Virelizier, L. Montagnier and U. Hazan, "A microtransfection method using the luciferase-encoding reporter gene for the assay of human immunodeficiency virus LTR promoter activity", *Gene 88*:197-205 1990.

20 Sherf, B. A., S. L. Navarro, R. R. Hannah and K. V. Wood, "Dual-luciferase™ reporter assay: an advanced co-reporter technology integrating firefly and <u>Renilla</u> luciferase assays. *Promega Notes 57*:2-9 (1996).

Stanley, P. E., "A review of the bioluminescent ATP techniques in rapid microbiology", *J. Biolum. Chemilum. 4*:375-380 (1989).

Tautoriainen, S., M. Virta, W. Chang and M. Karpi, "Measurement of firefly luciferase reporter gene activity from cells and lysates using Escherichia coli arsenite and mercury sensors", *Anal. Biochem. 272*:191-198 (1999).

Viviani, V. R., E. J. H. Bechara and Y. Ohmiya, "Cloning, sequence analysis, and expression of active Phrixothrix railroad-worm luciferases: Relationship between bioluminescence spectra and primary structures", *Biochemistry 38*:8271-8279 (1999).

Wood, K. V., "The chemical mechanism and evolutionary development of beetle bioluminescence", *Photochem. Photobiol. 62*:662-673 (1995).